

A many-valued logic for approximate reasoning

by Silvano Di Zenzo

A new system of many-valued logic, the Extended Post system of order p , $p \geq 2$, is proposed as a system of logic supporting reasoning with facts and rules which are reliable to a specified extent. In an Extended Post system there are as many operations of logical disjunction and logical conjunction as there are truth values. The truth value associated with a particular operation of disjunction (conjunction) acts as a threshold value controlling the behavior of the operation. The availability of an extended set of logical operations provides improved flexibility in the symbolic translation of sentences from the ordinary word-language. Extended Post systems are equipped with a semantics in which graded rather than crisp sets correspond to predicates. The system exhibits a "rich" algebraic structure. The p operations of disjunction form a distributivity cycle. To each disjunction there corresponds a dual operation of conjunction, the two operations being distributive to one another. The p conjunctions form a dual distributivity cycle. Both propositional calculus and first-order predicate calculus of EP systems are developed. The application to approximate reasoning is described. It is shown that there exist distinct isomorphic copies of fuzzy logic, each corresponding to a distinct level of approximation and being complete to resolution.

©Copyright 1988 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

1. Introduction

A knowledge-based system is able to perform one or more tasks relying upon its knowledge base as its primary source of information. Question-answering is one of the most usual tasks, and querying a knowledge base is one of the classical artificial intelligence applications.

In question-answering systems, facts and information are stored as axioms in some logical language, and the problem to be solved, or the question to be answered, is stated as a theorem to be proved (or refused) by means of the inference rules available in the language.

When the logical language is *complete*, all the logical consequences of the axioms are derivable as theorems. Then question-answering can be formulated as a theorem-proving task. This is the case when the logical language of first-order logic is adopted. However, first-order logic is not the only possible choice. Other logics can be adopted, for example to achieve a higher degree of expressiveness, or to handle temporal relationships.

The approach outlined above suffers from a drawback, which is that the facts and the rules taken as axioms are assumed to be absolutely correct. There is no room for uncertainty or approximation. On the other hand, it is well known that many of the rules provided by experts, and much of the data provided by users, are provided as being "true, but for a few possible exceptions," or "true in the majority of cases," or even "more or less true, where the degree of truth varies from case to case." In many applications the approximate nature of the notion of truth is intrinsic, and it would be highly desirable to have machines that can make intelligent decisions by relying both on incomplete knowledge and on rules of behavior that are reliable only to some specified extent.

Since logic is technically very well suited for making inferences and supporting reasoning, artificial intelligence

researchers have investigated various possible generalizations of logic for dealing with facts and rules which are only partially, or approximately, true. There has been interest in fuzzy logic, and in techniques of approximate reasoning related to the notion of a fuzzy set. There have also been efforts to combine logic with probability theory. Expert systems embodying reasoning methods based on Bayes' rule or "likelihood" rules have already been developed, and research aimed at providing theoretical bases for the *ad hoc* techniques is being done. Nilsson [1] presents a useful discussion of the subject and reviews many expert systems that rely on uncertain knowledge.

Shafer's theory of evidential reasoning [2] has recently received attention as a possible model for probabilistic reasoning in expert-system applications, and work has been done [3] to overcome the difficulties of implementing Shafer's belief functions in rule-based systems.

The purpose of this article is to provide an overview of a new system of many-valued logic, the Extended Post (*EP*) system. Some of the technical aspects of this system have already been discussed in the literature [4], with emphasis on the underlying algebra (which is a multiple Boolean algebra [5]). This paper, while self-contained, focuses on the semantics and the inference subsystem of the *EP* logic system. While problems encountered in the actual implementation of the resolution by *EP* logic may well need examination, they are outside the scope of the present discussion. We will show, however, that the system is complete to resolution and therefore that implementation in rule-based systems (specific cases of which would clearly need to be tested) is feasible.

2. Background and heuristic considerations

A sizable portion of the literature of logic does not adhere to the viewpoint that there are only two truth values, namely *truth* and *falsehood*.

Many-valued logics go back to ancient times. Thus Aristotelian logic recognized four truth modes, namely *necessity*, *contingency*, *possibility*, and *impossibility*.

In modern times, Lukasiewicz [6] first introduced a three-valued propositional calculus in 1920. He considered a third truth value, intermediate between *truth* and *falsehood*, expressing *possibility*. Denoting *truth* by 1, *falsehood* by 0, and the intermediate value by $\frac{1}{2}$, Lukasiewicz's truth tables for $\neg$ (negation), $\wedge$ (logical conjunction), $\vee$ (logical disjunction) can be summarized as follows:

$$T(\neg A) = 1 - T(A), \quad (1)$$

$$T(A \wedge B) = \min\{T(A), T(B)\}, \quad (2)$$

$$T(A \vee B) = \max\{T(A), T(B)\}, \quad (3)$$

where A, B are propositions, and $T(A)$ denotes the truth value of A .

Generalization to a many-valued propositional calculus using any number p of truth values, $p \geq 2$, was made by Post [7] in 1921, independently of Lukasiewicz. In a p -valued Post system there are p truth values $0, 1, \dots, p-1$, and the operations of logical conjunction and disjunction are defined by Equations (2) and (3). However, a different unary operation is used to express negation, namely,

$$T(\neg A) = T(A) + 1 \quad \text{for } T(A) < p-1,$$

$$T(\neg A) = 0 \quad \text{otherwise.}$$

Post's negation is not self-evident; however, its important advantage over Lukasiewicz's negation is that it leads to a truth-functionally complete system, i.e., a system in which every logical condition is expressible in terms of the connectives of the system.

After the pioneering work of Lukasiewicz and Post, numerous logicians have contributed to many-valued logic, and the field has become one of the most investigated subjects in logic.

The interest in many-valued logic is motivated by various application issues, such as supporting inexact and approximate reasoning and the design of multiple-valued switching circuits. Indeed, as two-valued propositional calculus is the basis for the design of two-valued logic circuits, so is many-valued propositional calculus the basis for the logical design of switching circuits with more than two stable states.

The size of the literature on many-valued logic is very large. For example, books on many-valued logic have been authored by Rosser and Turquette [8], Ackermann [9], and Rescher [10]. The book by Rescher contains a bibliography that is complete up to 1965. A review paper by Wolf [11] gives references updated to 1974. More recent reviews include Hurst [12] and Guccione et al. [13].

Since Zadeh introduced the notion of a fuzzy set, there has been an impressive growth of interest in *fuzzy logic*, and in techniques of approximate, imprecise, and inexact reasoning related to the notion of fuzzy sets.

Since we refer to both fuzzy sets and fuzzy logic, we shall recall some very basic notions on both subjects.

Let X be a nonempty set, and let I denote the unit interval of the real line, $I = [0, 1]$. After Zadeh [14], a fuzzy subset of X (equivalently, a fuzzy set in the universe X) is defined to be a function $f: X \rightarrow I$. Given any x in X , the degree of membership of x in X is $f(x)$. The union and the intersection of any two fuzzy sets f, g in the same universe X are defined as

$$(f \cup g)(x) = \max\{f(x), g(x)\},$$

$$(f \cap g)(x) = \min\{f(x), g(x)\}.$$

Equipped with union and intersection, the set $\mathcal{L}(X)$ of the fuzzy subsets of X is a complete distributive lattice. The least element of $\mathcal{L}(X)$ is the function $X \rightarrow I$ which is everywhere 0; the greatest element is the function which is everywhere 1.

The complement $\bar{f}$ of a fuzzy set f is defined by $\bar{f}(x) = 1 - f(x)$.

Equipped with union, intersection, and complement, $\mathcal{L}(X)$ is a deMorgan algebra [22] (with complement acting as involution).

A more specialized notion of fuzzy sets, that of *graded sets*, is obtained if the unit interval I is replaced by any finite subset of I containing both 0 and 1. Graded sets exhibit additional algebraic properties that are useful in many applications [5]. For example, graded subsets of the universe of discourse are assigned to monadic predicates by interpretations in many-valued logic.

The term *fuzzy logic* does not have a unique meaning in the literature. Sometimes it addresses a *local logic* in which the truth values are fuzzy subsets of the unit interval [15]. Other variants of fuzzy logic and their possible applications in expert systems are discussed in Zadeh [16] and Prade [17]. The latter paper also contains extended references to recent literature on the subject.

In the present paper, by *fuzzy logic* we mean the special many-valued logic investigated by Marinos [18], Chang and Lee [19], Lee [20], and Aronson et al. [21], among others. In this sense, fuzzy logic is a many-valued logic in which the truth-set is the unit interval of the real line, and the logical operations of negation, conjunction, and disjunction are those expressed by (1)–(3). In this restricted sense, fuzzy logic is a straight generalization of Lukasiewicz's logic.

We conclude our extremely concise account of many-valued logic by quoting a note of criticism made by Birkhoff on, in a sense, all of the efforts to construct many-valued logics [22]. The core of Birkhoff's argument is the following: As long as we equip the set of all the propositions with one logical conjunction and one logical disjunction, we are implicitly treating that set as a Boolean algebra, and there is no structure-preserving application of a Boolean algebra to a linearly ordered set with more than two elements.

We shall keep this remark in mind when defining our system of many-valued logic. Actually, we do not assign a structure of Boolean algebra to the set of all propositions.

Our starting point is our everyday language, and the way we use it in everyday speech. Of course, we use language not only for communication but also for reasoning. There are both analogies and differences between the way we use a nonformal language like English for reasoning about our everyday affairs, and the way a mathematician uses it for proving a theorem. In both cases arguments such as "if A is true, then B is true" or, say, "if either A or B is true, then C is false" are used. However, in everyday speech we are less drastic. For example, when we use arguments of the form "if A , then B ," we usually mean " B is likely to be true if A is true," or " B is often true when A is true," and so forth.

Even the statement that A is true may have different meanings according to the context in which it is made.

We concentrate on certain patterns of informal reasoning, as performed in nonformal language, which seem to suggest that certain more articulate logical operations might be useful. In current speech we often refer to properties and relations that are obviously complex and require compound predicates for adequate translation into a symbolic language. The resulting compound predicates are often assembled from simpler components by means of certain repetitive patterns of *and* and *or*.

Let us consider some more examples. Suppose that b is a person, and we assert that b is a computer scientist. If we analyze the meaning of this statement, we easily recognize that it is a very complex statement, in which numerous assertions about b are implied. When we say that someone is a computer scientist, we usually understand that he or she knows something of all the relevant subfields of computer science, and, very likely, knows almost everything of some of these subfields. Thus, "computer scientist" is indeed a composite property; in its definition, a certain pattern of logical conjunctions and logical disjunctions occurs.

This pattern repeats itself at various levels. For example, let us make a "zoom" over two of the relevant subfields of "computer science," namely "programming languages" and "artificial intelligence." When we say that someone has some knowledge of programming languages, we probably mean that he has heard of all the relevant programming languages, and perhaps has good knowledge of a few. When we say that somebody is a researcher in artificial intelligence, we probably mean that he understands the basic facts of, say, high- and low-level vision, robotics, knowledge representation, etc., and has deep knowledge of, say, theorem proving and automated reasoning (or vice versa, or any other combination of well-known and less-known topics in artificial intelligence).

In everyday language we also combine these complex (and inexact) properties and relations to form more complex and structured attributes. Again, the new attributes do not often fit such schemes as pure logical conjunction or disjunction. Often they occur as compound notions sharing features from both conjunction and disjunction.

Analogous situations are found in various other semantic domains; for example, "beautiful," "clever," "polite" are complex attributes, and their symbolic translation into a formal language may be easier if certain weak, or generalized, logical operations of conjunction and disjunction are available. The same generalized operations may be useful for assembling new, and more complex, properties from "beautiful," "clever," "polite," etc.

In this paper we introduce logical operations which, in a sense, mimic this mixed and graded character of the connectives actually used in informal reasoning. We introduce two families of generalized logical operations, a family of generalized disjunctions and a family of generalized conjunctions.

We define a symbolic translation for those statements of the conversational language which assert that the facts $A, B, \dots, C$ are all more or less true and that possibly some of them are quite true. For each grade of truth m we define a generalized (or weak) operation of logical conjunction $\bigwedge_m$ such that the proposition

$$A \bigwedge_m B \bigwedge_m \dots \bigwedge_m C$$

has a degree of truth greater than or equal to m if and only if all the facts $A, B, \dots, C$ have a degree of truth of at least m . If this condition is met, the truth value of the above proposition will be set to be equal to the maximum of the truth values $T(A), T(B), \dots, T(C)$.

A dual operation of logical disjunction corresponding to $\bigwedge_m$ will be defined on the basis of purely algebraic considerations. For each grade of truth m we shall define a generalized (or weak) operation of logical disjunction $\bigvee_m$ such that the proposition

$$A \bigvee_m B \bigvee_m \dots \bigvee_m C$$

has a degree of truth greater than or equal to m if and only if at least one of the facts $A, B, \dots, C$ has a degree of truth of at least m . If this condition is met, the truth value of the proposition is set equal to the minimum of those values $T(A), T(B), \dots, T(C)$ which are above the threshold m .

3. Extended Post systems

Any logical system has three basic components: a formal language, a semantics, and a set of inference rules.

The language of a logical system is a set of well-formed formulas built from certain primitive symbols according to given rules of formation. So long as we are concerned with the syntactical definition of the language, the formulas of the language remain uninterpreted.

The language of a logical system is so constructed that various sentences of the natural language can be translated into it, provided only that certain signs have received determinate interpretations. Interpretations are dealt with in the semantics of the language.

The rules of inference permit us to deduce statements in the language from other statements in the language.

We specify a logical system, the Extended Post system of order p [or the p -valued EP system, or even the EP(p) system, for short], where p is any integer greater than or equal to 2. In this section we define the language of the system and its semantics, while in the next section we present and discuss the inference rules. For $p = 2$, the p -valued EP system reduces to ordinary two-valued first-order logic.

It may be possible to define the language of the p -valued EP system along purely syntactical lines, without mentioning semantic notions such as *truth* and *falsehood* and interpretations. It is, however, more convenient to take into

account, from the very beginning, that our formal language is to be used as a medium for expressing statements from the conversational language. A great many of these statements (some would say all) are about some specific domain of objects; i.e., they describe *some* of the properties of certain objects and *some* of the relations among them. Thus, we find it convenient to refer from the outset to a fixed, though unspecified, domain D of individual objects. And we shall bear in mind that the well-formed formulas of our language must be suitable for interpretation as assertions concerning the objects in D .

We begin by discussing the semantic notion of truth-set. The truth-set is the set of the possible truth values of a proposition. In two-valued logic, *truth* and *falsehood* are the only possible truth values of a proposition. If *truth* is designated by 1 and *falsehood* by 0, $V(2) = \{0, 1\}$ is the truth-set of two-valued logic. In a p -valued EP system, the truth-set is defined to be

$$V(p) = \{0, 1/(p-1), \dots, (p-2)/(p-1), 1\}.$$

A proposition can take any element of $V(p)$ as a truth value. We regard the elements of the set $V(p)$ as actually representing various degrees of truth. For example, the truth values in a four-valued system might be interpreted as *truth*, *plausibility*, *implausibility*, and *falsehood*.

For notational convenience, each truth value is assigned a nonnegative integer as its label. To this end, let $I(p)$ denote the set of the first p nonnegative integers, $I(p) = \{0, 1, \dots, p-1\}$, and let $u: I(p) \rightarrow V(p)$ be such that $u(m) = m/(p-1)$ for all m in $I(p)$. Any m in $I(p)$ may then be regarded as a pointer to the truth value $u(m)$, or, equivalently, as the label of $u(m)$.

The set $I(p)$ possesses a number of algebraic properties. Equipped with $\vee$ (maximum) and $\wedge$ (minimum), $I(p)$ is a complete distributive lattice. The least element is 0; the greatest is $p-1$. Equipped with $\vee$, $\wedge$, and $\sim$, where $\sim$ is the unary operation such that $\sim m = p-1-m$, $I(p)$ is a deMorgan algebra. Equipped with $\vee$, $\wedge$, and $[\cdot]$, where $[\cdot]$ is the unary operation (called *cycle*) such that $[m] = m+1$ for $m \neq p-1$ and $[p-1] = 0$, $I(p)$ is a Post algebra of order p . By the bijection u , the algebraic structures defined on $I(p)$ can be carried over onto $V(p)$. In particular, for any v in $V(p)$, given that $v = u(m)$, we define $[v]$ and $\sim v$ by means of $[v] = u[m]$ and $\sim v = u(\sim m)$, respectively. Note that $\sim v$ turns out to be just $1-v$.

The primitive symbols of our languages are (1) parentheses, (2) individual, functional, and predicate symbols, and (3) logical connectives and quantifiers.

There are two parentheses, the left parenthesis "(" and the right parenthesis ")".

The individual symbols are names, or (if one prefers) designations, for individual objects. There are two kinds of individual symbols, namely individual constants and individual variables. An individual constant is a designation

of a specific member of D ; an individual variable is a designation of a generic member of D . For individual constants we use lowercase letters $a, b, c, \dots$ taken from the beginning of the alphabet. For individual variables we use lowercase letters $x, y, z, \dots$ taken from the end of the alphabet.

An n -place function symbol is a designation of a function on D^n to D , where D^n is the n th Cartesian power of D . For function symbols we use lowercase letters such as $f, g, h, \dots$ taken from the middle of the alphabet. In examples, we also use for function symbols various letter groups, all lowercase. For example, if the domain of individuals D is the set of the natural numbers, then *prod* might designate the product of two such numbers (a function from $D \times D$ to D). Thus, if a and b are individual constants, hence designations of certain numbers, then *prod*(a, b) denotes the product of these numbers.

An n -place predicate symbol is a designation of a function on D^n to the truth-set $V(p)$. Note that, for $p = 2$, an n -place predicate symbol turns out to be a designation of a subset of D^n , i.e., a name for an n -place relation among the individuals in D . For $p > 2$, an n -place predicate symbol is a designation of a graded subset of D^n , i.e., an n -place graded relation in D . For predicate symbols we use the capital letters $P, Q, R, \dots$. In connection with examples, we also use for predicates significative letter groups with the first letter capitalized.

Terms and atomic formulas are defined as in two-valued logic.

Terms are defined recursively as follows: (1) An individual constant is a term. (2) An individual variable is a term. (3) If f is an n -place function symbol and $t_1, \dots, t_n$ are terms, then $f(t_1, \dots, t_n)$ is a term. (4) Nothing is a term unless its being so follows from (1)–(3).

If P is an n -place predicate symbol and $t_1, \dots, t_n$ are terms, then $P(t_1, \dots, t_n)$ is an atomic formula (briefly, an atom). Nothing else is an atomic formula.

Example If our language were to be applied to a domain comprising a specified group of people, a might be taken as an abbreviation for *Charles Smith*, b for *Mary Miller*, etc. In this domain *Beautiful* might designate the property of being beautiful, and *Likes* might designate the relation that obtains between two persons when the first likes the second. Both this property and this relation are graded attributes: Beauty is distributed among people in various degrees, and there are various grades of appreciation between any two persons. According to the above definition, *Beautiful*(b) and *Likes*(a, b) are atomic formulas in our language. *Beautiful*(b) is the translation into our formal language of the English sentence "Mary Miller is beautiful," while *Likes*(a, b) is the translation of "Charles Smith likes Mary Miller."

In a preceding section we gave heuristic motivations for the association of distinct logical operations of disjunction

and conjunction with each truth value. We now list the logical connectives of a p -valued EP system, and specify formally what strings of signs are to be considered well-formed formulas in the language of the system.

The logical connectives of the language are $\sim$ (not), $[.]$ (cycle), $\bigvee_m$ (m -disjunction), $\bigwedge_m$ (m -conjunction), $\supset_m$ (m -implication), and $\equiv_m$ (m -equivalence), where m is any element of $I(p)$.

We save parentheses by assigning decreasing ranks to $\equiv_m, \supset_m, \bigvee_m, \bigwedge_m, \sim, [.]$, and requiring that the connective with greater rank reach further.

Well-formed formulas (WFFs) are defined recursively by the following rules of formation:

RF1. Atoms are WFFs.

RF2. If A is a WFF, then $\sim A$ and $[A]$ are WFFs.

RF3. If A and B are WFFs, then $A \equiv_m B$, $A \supset_m B$, $A \bigvee_m B$, $A \bigwedge_m B$ are WFFs.

RF4. If F is a WFF, and x is a free variable in F , then both $(\bigvee_m x)F$ and $(\bigwedge_m x)F$ are WFFs.

It is understood that m can be any member of $I(p)$. With reference to Rule RF4, bound and free variables are here defined in exactly the same way as in ordinary two-valued logic.

With the definition of well-formed formulas, the specification of the syntax of the language is completed. In connection with the specification of the language, a few aspects of the semantics have already been mentioned. However, the specification of the semantics is not completed. It remains for us to specify what we mean by an interpretation, and how the truth value of a formula is evaluated under a given interpretation.

To define an interpretation for a formula we must specify a domain D together with an assignment of values to all value-bearing signs in the formula. More specifically, an interpretation I of a formula F consists of a nonempty set D , and an assignment of values to each constant, function symbol, and predicate symbol occurring in F as follows:

- To each individual constant, assign an element in D .
- To each n -place function symbol, assign a mapping $D^n \rightarrow D$.
- To each n -place predicate symbol, assign a mapping $D^n \rightarrow V(p)$.

Note that, at least in general, an interpretation assigns a graded subset of D^n to each n -place predicate symbol.

Let $T(A)$ (TA , if there is no danger of confusion) denote the truth value of a well-formed formula A . For every interpretation of a formula over a domain D , the formula can be evaluated to one of the truth values in $V(p)$ according to the following rules of evaluation:

$$\text{RE1. } T(\sim A) = 1 - T(A).$$

$$\text{RE2. } T[A] = [TA].$$

$$\text{RE3. } T(A \underset{m}{\vee} B) = \begin{cases} \min\{TA, TB\} & \text{if } TA, TB < u(m) \\ & \text{or } TA, TB \geq u(m), \\ \max\{TA, TB\} & \text{otherwise.} \end{cases}$$

$$\text{RE4. } T(A \underset{m}{\wedge} B) = \begin{cases} \max\{TA, TB\} & \text{if } TA, TB < u(m) \\ & \text{or } TA, TB \geq u(m), \\ \min\{TA, TB\} & \text{otherwise.} \end{cases}$$

$$\text{RE5. } T(A \supset B) = T((\sim A) \underset{m}{\vee} B).$$

$$\text{RE6. } T(A \equiv B) = T((A \supset B) \underset{m}{\wedge} (B \supset A)).$$

$$\text{RE7. } T((\underset{m}{\vee} x)F) = \inf\{T(F(d)) : d \in D, T(F(d)) \geq u(m)\} \\ \vee \inf\{T(F(d)) : d \in D, T(F(d)) < u(m)\}.$$

$$\text{RE8. } T((\underset{m}{\wedge} x)F) = \sup\{T(F(d)) : d \in D, T(F(d)) \geq u(m)\} \\ \wedge \sup\{T(F(d)) : d \in D, T(F(d)) < u(m)\}.$$

A few comments may be convenient. Rule RE1 is the rule of evaluation for negation in fuzzy logic [19, 20], while RE2 is the analogous rule in the many-valued propositional calculus introduced by Post [7]. Rules RE3 and RE4 are the explications here proposed for the weak (or generalized) notions of logical disjunction and conjunction that have been informally discussed in the preceding section. Rules RE5 and RE6 define the truth values of $A \supset B$ and $A \equiv B$ in terms of those of $\sim A$, $A \underset{m}{\vee} B$, and $A \underset{m}{\wedge} B$. Indeed, $A \supset B$ and $A \equiv B$ are to be regarded as abbreviations of $\sim A \underset{m}{\vee} B$ and, respectively, $(A \supset B) \underset{m}{\wedge} (B \supset A)$. Finally, RE7 and RE8 are the infinitary versions of RE3 and RE4, respectively.

For the sake of clarity, we rephrase RE8 in words. Let α be the maximum of the values $T(F(d))$ computed over those $d \in D$ for which $T(F(d)) \geq u(m)$; let β be the maximum of the values $T(F(d))$ computed over those for $d \in D$ for which $T(F(d)) < u(m)$. Then the truth value of the formula $(\underset{m}{\wedge} x)F$ is $\min(\alpha, \beta)$. It is understood that the least upper bound and the greatest lower bound of an empty set of members of $V(p)$ are 1 and 0, respectively.

Example Let P be a monadic (one-place) predicate symbol. Consider the formula $(\underset{3}{\wedge} x)P(x)$ in a six-valued EP system. Assume that the domain is $D = \{1, 2\}$.

1. Consider an interpretation in which the assignment for P is $P(1) = 1/5 = 0.2$, $P(2) = 4/5 = 0.8$. In this case

$$\{d \in D : T(P(d)) \geq u(3)\} \\ = \{d \in D : T(P(d)) \geq 0.6\} = \{2\},$$

$$\{d \in D : T(P(d)) < 0.6\} = \{1\}.$$

$$\text{Thus, } \alpha = 0.8, \beta = 0.2, T((\underset{3}{\wedge} x)P(x)) = \min(\alpha, \beta) = 0.2.$$

$\hat{\wedge}$	0	$\frac{1}{2}$	1
0	0	$\frac{1}{2}$	1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
1	1	1	1

$\hat{\vee}$	0	$\frac{1}{2}$	1
0	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	1
1	0	1	1

$\hat{\supset}$	0	$\frac{1}{2}$	1
0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	1

Figure 1

Truth tables of the operations of logical conjunction and disjunction of the three-valued Extended Post system.

2. Consider an interpretation in which the assignment for P is $P(1) = 1/5 = 0.2$, $P(2) = 0$. In this case

$$\{d \in D : T(P(d)) \geq 0.6\} = \emptyset,$$

$$\{d \in D : T(P(d)) < 0.6\} = D.$$

$$\text{Thus, } \alpha = 1, \beta = 0.2, T((\underset{3}{\wedge} x)P(x)) = \min(\alpha, \beta) = 0.2.$$

The truth tables of the operations of the $EP(3)$ system are given in Figure 1.

4. Algebraic properties of the Extended Post system

At this point both the syntax and the semantics of our system have been specified. A few comments may again be convenient. We have been led to our system by two basic issues. On one hand, we have tried to mimic everyday conversational language by introducing a set of new logical operations each sharing features of both the conjunction and the disjunction of usual logic. On the other hand, we have tried to meet the formal requirements set forth by Birkhoff for many-valued logics. We are now interested in investigating the formal properties of our system. In particular, we are interested in understanding to what extent the algebra which underlies the system can be regarded as actually generalizing Boolean algebra, i.e., the algebra of ordinary two-valued logic.

Our first step is to establish a lemma showing that there exists a bijection $I(p) \rightarrow I(p)$, namely the cycle operation, which carries $\underset{m}{\vee}$ onto $\underset{[m]}{\vee}$ and $\underset{m}{\wedge}$ onto $\underset{[m]}{\wedge}$. As an immediate consequence, all of the operations of m -disjunctions are pairwise isomorphic: In a sense, any one of them is a shifted copy of any other. The same applies to m -conjunctions. The

lemma will prove useful since it allows us to transfer everything we know of the pair of operations $(\bigvee_0, \bigwedge_0)$ to any pair of operations $(\bigvee_m, \bigwedge_m)$. It follows from RE3 and RE4 that $\bigvee_0, \bigwedge_0$ coincide with the well-known operations of minimum $\wedge$ and maximum $\vee$, respectively, in the truth-set $V(p)$; hence we have detailed knowledge of them.

We shall say that two formulas A and B are *logically equivalent* if A and B always take the same truth value in any interpretation. In some cases we use the sign $\approx$ of the metalanguage as an abbreviation of "is logically equivalent to."

Lemma For any $m \in I(p)$ and any pair A, B of ground atomic formulas (i.e., atomic formulas not containing variables),

$$(a) [A \bigvee_m B] \approx [A] \bigvee_m [B],$$

$$(b) [A \bigwedge_m B] \approx [A] \bigwedge_m [B].$$

Proof Only part (a) is proved; part (b) can be proved analogously. Let $m < p - 1$. We may have $TA \geq u(m + 1)$, or $TA = u(m)$, or $TA < u(m)$, and independently, $TB \geq u(m + 1)$, or $TB = u(m)$, or $TB < u(m)$. Thus, nine cases should be considered. Since $\bigvee_m$ is obviously commutative, we consider only six cases:

1. $TA, TB \geq u(m + 1)$. Then also $TA, TB \geq u(m)$, hence $T(A \bigvee_m B) = TA \wedge TB$, whence $T[A \bigvee_m B] = [TA \wedge TB]$. On the other hand, $T[A], T[B]$ are either greater than $u(m + 1)$ or 0:
 - (i) If they are both greater than $u(m + 1)$ or both 0, then $T([A] \bigvee_m [B]) = T[A] \wedge T[B] = [TA] \wedge [TB] = [TA \wedge TB]$.
 - (ii) If, say, $T[A] = 0$ and $T[B] \geq u(m + 1)$, $T([A] \bigvee_m [B]) = 0 \vee T[B] = T[B] = [TA \wedge TB]$, since, in this case, $TA = 1$.
2. $TA \geq u(m + 1)$, $TB = u(m)$. In this case $T(A \bigvee_m B) = u(m)$ by RE3, hence $T[A \bigvee_m B] = u[m]$. On the other hand, $T[B] = u[m]$, hence, again by RE3, $T([A] \bigvee_m [B]) = u[m]$.
3. $TA \geq u(m + 1)$, $TB < u(m)$. $T(A \bigvee_m B) = TA \vee TB = TA$ thus $T[A \bigvee_m B] = [TA]$. On the other hand, if $T[A] = 0$, then $T([A] \bigvee_m [B]) = 0 = [TA]$. If $T[A] \geq u(m + 1)$, then $T([A] \bigvee_m [B]) = T[A] \vee T[B] = T[A] = [TA]$.
4. $TA = TB = u(m)$. In this case the proposition holds trivially.
5. $TA = u(m)$, $TB < u(m)$. $T[A \bigvee_m B] = T([A] \bigvee_m [B]) = u[m]$ by RE3.
6. $TA, TB < u(m)$. $T(A \bigvee_m B) = TA \wedge TB$, thus $T[A \bigvee_m B] = [TA \wedge TB]$. On the other hand, $T[A], T[B] < u(m + 1)$, thus $T([A] \bigvee_m [B]) = T[A] \wedge T[B] = [TA \wedge TB]$.

It is left to the reader to check that the proposition holds for $m = p - 1$.

As noted, we have detailed knowledge of the properties of $\bigvee_0$ and $\bigwedge_0$. Both operations are associative, commutative, idempotent, and equipped with both neutral and absorbing elements. Indeed, if B is a 0-tautology, i.e., if it takes the truth value 0 in all interpretations, then B is absorbing for $\bigvee_0$ and neutral for $\bigwedge_0$; i.e., for any WFF A ,

$$A \bigvee_0 B \approx B, \quad A \bigwedge_0 B \approx A.$$

If C is a 1-tautology, C is neutral for $\bigvee_0$ and absorbing for $\bigwedge_0$, or

$$A \bigvee_0 C \approx A, \quad A \bigwedge_0 C \approx C.$$

The two operations $\bigvee_0$ and $\bigwedge_0$ are distributive to one another. Besides, there exists a unary operation $\sim$ such that

$$\sim \sim A \approx A,$$

$$\sim(A \bigvee_0 B) \approx (\sim A) \bigwedge_0 (\sim B),$$

$$\sim(A \bigwedge_0 B) \approx (\sim A) \bigvee_0 (\sim B).$$

By the above lemma, all these properties can be transferred from the pair of operations $(\bigvee_0, \bigwedge_0)$ to any pair $(\bigvee_m, \bigwedge_m)$. Thus, we have the following propositions, in which m is any element of $I(p)$, and A, B, C are any well-formed formulas.

Proposition 1

$$(a) (A \bigvee_m B) \bigvee_m C \approx A \bigvee_m (B \bigvee_m C),$$

$$(A \bigwedge_m B) \bigwedge_m C \approx A \bigwedge_m (B \bigwedge_m C),$$

$$(b) A \bigvee_m B \approx B \bigvee_m A, \quad A \bigwedge_m B \approx B \bigwedge_m A,$$

$$(c) A \bigvee_m A \approx A, \quad A \bigwedge_m A \approx A.$$

Proposition 2

$$(a) A \bigvee_m (B \bigwedge_m C) \approx (A \bigvee_m B) \bigwedge_m (A \bigvee_m C),$$

$$(b) A \bigwedge_m (B \bigvee_m C) \approx (A \bigwedge_m B) \bigvee_m (A \bigwedge_m C).$$

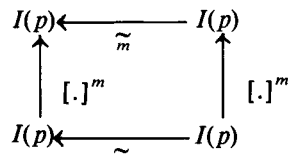
Let n be the inverse image of m in the cycle operation; i.e., $[n] = m$. Let M and N be an m -tautology and an n -tautology, respectively.

Proposition 3

$$(a) A \bigvee_m M \approx M, \quad A \bigwedge_m M \approx A,$$

$$(b) A \bigvee_m N \approx A, \quad A \bigwedge_m N \approx N.$$

Let us denote by $\approx_m$ the mapping $I(p) \rightarrow I(p)$ which implies the following commutative diagram:



Considered as a unary operation in the set of all well-formed formulas, the operation $\tilde{m}$ is such that, for any WFF A ,

$$\tilde{m}A \approx [\sim[A]^{-m}]^m.$$

Since $\sim$ is the involution of the deMorgan algebra associated with the pair $(\bigvee, \bigwedge)$, $\tilde{m}$ is the involution of the deMorgan algebra associated with the pair of operations $(\bigvee_m, \bigwedge_m)$.

Proposition 4

- (a) $\tilde{m} \tilde{m}A \approx A$,
- (b) $\tilde{m}(A \bigvee_m B) \approx (\tilde{m}A) \bigwedge_m (\tilde{m}B)$,
- (c) $\tilde{m}(A \bigwedge_m B) \approx (\tilde{m}A) \bigvee_m (\tilde{m}B)$.

For all m in $I(p)$, the set of all WFFs equipped with $\bigvee_m, \bigwedge_m$, and $\tilde{m}$ is a deMorgan algebra. We now investigate whether there are any relationships among all these algebras.

The method of proceeding based on pure semantics, i.e., rules RE1-RE8, which we have followed up to now, is the most usual in logic. For the results to be derived next, however, it would lead to laborious proofs. For the sake of brevity, in the rest of this section we proceed along algebraic lines.

We define the $2p$ binary operations $\bigvee, \dots, \bigvee_{p-1}, \bigwedge, \dots, \bigwedge_{p-1}$ over the set $I(p)$,

$$x \bigvee_m y = \begin{cases} \min(x, y) & \text{if } x, y < m \text{ or } x, y \geq m, \\ \max(x, y) & \text{otherwise,} \end{cases}$$

$$x \bigwedge_m y = \begin{cases} \max(x, y) & \text{if } x, y < m \text{ or } x, y \geq m, \\ \min(x, y) & \text{otherwise.} \end{cases}$$

Equipped with these operations and the unary operations $\sim$ and $[.]$, the set $I(p)$ is denoted $\mathcal{L}(p)$. Obviously, everything said of $\bigvee_m$ and $\bigwedge_m$ as logical operations between WFFs of the $EP(p)$ system can be reformulated to apply to $\mathcal{L}(p)$. For example, the above lemma states that, for any x, y, m in $I(p)$, we have

$$[x \bigvee_m y] = [x] \bigvee_{[m]} [y],$$

$$[x \bigwedge_m y] = [x] \bigwedge_{[m]} [y].$$

The converse holds; that is, the results derived for $\mathcal{L}(p)$ can be reformulated to apply to the set of WFFs of a p -valued EP system.

Remark Everything said of $I(p)$ can be restated in terms of $V(p)$. It is understood that all the algebraic structures defined on $I(p)$ are carried over onto $V(p)$ along the bijection u .

Proposition 5

For all m in $I(p)$,

- (a) $\bigvee_m$ is distributive with respect to $\bigvee_{[m]}$,
- (b) $\bigwedge_m$ is distributive with respect to $\bigwedge_{[m]}$.

Proof Only part (a) is proved; part (b) can be proved analogously. $\bigvee$ is distributive to $\bigvee$ by direct verification. Now assume that $\bigvee_m$ is distributive to $\bigvee_{[m]}$. By repeated application of the basic lemma, we have

$$\begin{aligned} (x \bigvee_{[m]} y) \bigvee_{[m]} z &= [[x]^{-1} \bigvee_{[m]} [y]^{-1}] \bigvee_{[m]} [[z]^{-1}] \\ &= [[x]^{-1} \bigvee_{[m]} [y]^{-1}] \bigvee_m [z]^{-1} \\ &= [[x]^{-1} \bigvee_m [z]^{-1}] \bigvee_{[m]} ([y]^{-1} \bigvee_m [z]^{-1}) \\ &= [[x]^{-1} \bigvee_m [z]^{-1}] \bigvee_{[m]} [[y]^{-1} \bigvee_m [z]^{-1}] \\ &= (x \bigvee_m z) \bigvee_{[m]} (y \bigvee_m z). \end{aligned}$$

Proposition 6

For all x, m in $I(p)$,

$$x \bigvee_m [x] \bigvee_m \dots \bigvee_m [x]^{p-1} = m,$$

$$x \bigwedge_m [x] \bigwedge_m \dots \bigwedge_m [x]^{p-1} = [m]^{-1}.$$

Proof The elements $x, [x], \dots, [x]^{p-1}$ form the totality of the elements of $I(p)$. Since $\bigvee_m$ is associative, commutative, and idempotent, for any y in $I(p)$ we have

$$y \bigvee_m (x \bigvee_m [x] \bigvee_m \dots \bigvee_m [x]^{p-1}) = x \bigvee_m [x] \bigvee_m \dots \bigvee_m [x]^{p-1}.$$

Since there can be at most one absorbing element, the first equality is proved. The second equality can be proved analogously.

In Proposition 3 we mentioned m -tautologies. Proposition 6 proves the existence of m -tautologies for all m . Indeed, if A is any WFF, then

$$A \bigvee_m [A] \bigvee_m \dots \bigvee_m [A]^{p-1}$$

takes the truth value $u(m)$ in all interpretations.

Proposition 7

- (a) $\sim(x \bigvee_m y) = (\sim x) \bigwedge_{p-m} (\sim y)$,
- (b) $\sim(x \bigwedge_m y) = (\sim x) \bigvee_{p-m} (\sim y)$.

Proof Only part (a) is proved; part (b) can be proved analogously. Assume that either $x, y \geq m$ or $x, y < m$. Then $x \bigvee_m y = x \bigwedge y$. It follows that $\sim(x \bigvee_m y) = \sim(x \bigwedge y) =$

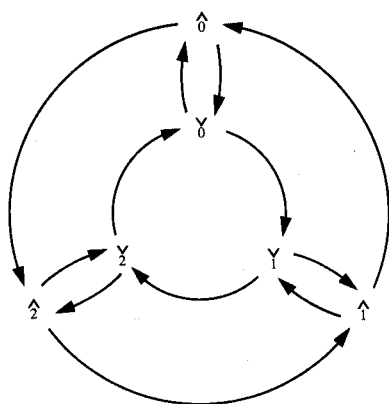


Figure 2

Distributivity graph of the operations of logical conjunction and disjunction of the Extended Post system of order 3.

$(\sim x) \vee (\sim y)$. On the other hand, our assumption on x, y is equivalent to the assumption that either $\sim x, \sim y < p - m$ or $\sim x, \sim y \geq p - m$. Thus $(\sim x) \wedge_{p-m} (\sim y) = (\sim x) \vee (\sim y)$. Now assume that $x \geq m, y < m$. Then $x \vee_m y = x \vee y$, whence $\sim(x \vee_m y) = \sim(x \vee y) = (\sim x) \wedge (\sim y)$. But $x \geq m$ implies $\sim x < p - m$, and $y < m$ implies $\sim y \geq p - m$. Thus $(\sim x) \wedge_{p-m} (\sim y) = (\sim x) \wedge (\sim y)$.

On the basis of the preceding propositions, we may try to draw a picture of a p -valued EP system. We have p logical operations of disjunction $\vee, \vee_1, \dots, \vee_{p-1}$, all exhibiting the usual properties of binary logical operations. Each operation is associative, commutative, idempotent, and equipped with both an identity and an absorbing element.

Taken together, these operations form a *distributivity cycle*; i.e., $\vee$ is distributive to $\vee_1, \vee_1$ is distributive to $\vee_2, \dots, \vee_{p-1}$ is distributive to $\vee$. To each disjunction $\vee_m$ there corresponds a dual operation of conjunction $\wedge_m$. Pairs of dual operations are distributive to one another; in addition, the identity element of one operation is absorbing for the dual operation. Together with $\sim_m$, each pair $(\vee_m, \wedge_m)$ forms a deMorgan algebra. The conjunctions exhibit the same algebraic properties as the disjunctions. Taken together, they form a distributivity cycle in the reverse order; i.e., $\wedge_0$ is distributive to $\wedge_{p-1}, \wedge_{p-1}$ is distributive to $\wedge_{p-2}, \dots, \wedge_1$ is distributive to $\wedge_0$.

Thus, a p -valued EP system exhibits an extremely uniform algebraic structure. This uniformity can be made even more evident by considering the symmetries, i.e., the

automorphisms, of the system. Let $G(p)$ be a graph such that (1) its vertices are the operations $\vee_m$ and $\wedge_m, m \in I(p)$, (2) there is an arc from vertex a to vertex b whenever a is distributive to b . For $p = 3$, $G(p)$ is shown in Figure 2. It can be verified easily that the group of the automorphisms of $G(p)$ is isomorphic to the dihedral group of order $2p$. That implies that, given any two operations a, b of disjunction (or conjunction), there is an automorphism which carries a into b . From a logical point of view, that implies that the truth values in $V(p)$ all stand on one and the same footing, and one can choose the pair 0, 1 in p distinct and logically equivalent ways.

So far, the algebraic properties of a p -valued EP system have been investigated to some extent. Now it may be of interest to relate the results derived above to the results obtained on Post algebras after Post's seminal work. The rest of this section is dedicated to a brief comparison of results.

The first algebra of p -valued truth-functionally complete logic corresponding to the work of Post was formulated by Rosenbloom [23] and called Post algebra of order $p, p \geq 2$. The currently adopted formulation of Post algebras using disjoint operators was given by Epstein [24]. Epstein improved Rosenbloom's axiom set considerably, and gave a complete representation theory of Post algebras.

One of Epstein's most important results was that every Post algebra of order p is isomorphic to a lattice of continuous functions from a Boolean space to a discrete p -element chain. This result leads to formulas that formally resemble those of a p -valued EP system. For example, when the class of Post algebras of order p is described as an equational class of algebras, p unary operations $C_0, C_1, \dots, C_{p-1}$ are introduced, satisfying

$$C_m(x \vee y) = C_m(x) \vee C_m(y), \quad (4)$$

$$C_m(x \wedge y) = C_m(x) \wedge C_m(y). \quad (5)$$

These unary operations resemble the operations $\sim_m$ that occur in Proposition 4. However, the two sets of unary operations are different, and there is no simple way of expressing one in terms of the other. For example, the C_m satisfy

$$C_m(x) \wedge \sim C_m(x) = C_m(x) \wedge \sim C_m(x) = 1,$$

while the operations $\sim_m$ do not.

After the work of Epstein, numerous authors focused their attention on generalizations of Post algebras. The review paper by Dwinger [25] provides a detailed account of the relevant efforts in this direction. We mention the generalized Post algebras of order ω^+ since they also are based on (infinitely many) unary operations satisfying Equations (4) and (5) but having yet a different meaning.

Generalized Post algebras of order ω^+ have been studied by H. Rasiowa [26, 27] and E. Orłowska [28, 29], among others.

By way of conclusion, we note that in Epstein's axiomatics for Post algebras, and in all the numerous generalizations of

Post algebras proposed so far, only the two binary lattice operations $\vee$ and $\wedge$ occur. Thus, the adoption (after proof of existence) of distributive cycles of generalized ANDs and ORs is unique to EP systems.

5. Inference in Extended Post systems

Up to this point we have considered the syntactic and the semantical aspects of EP systems. In this section we investigate inference in EP systems. We also discuss possible applications of EP systems in problem-solving and question-answering systems.

We assume that the available knowledge base consists of facts and rules which are reliable to some specified degree of confidence. They are not necessarily *fuzzy* or *vague*. What is relevant is the effect upon them of an error whose maximum magnitude is assumed to be known.

We consider the problem of deriving consequences from the knowledge base and, at the same time, of associating a degree of truth with each of the consequences that we derive.

We begin by considering a very simple case. Assume that we work with, say, $p = 100$ truth values, and all the data we start with are affected by the same maximum error of, say, 20 percent. In this situation we might be interested in logical consequences of our data within a degree of confidence of 80 percent.

Note that this is a simplified situation. In a realistic situation we start with facts and rules each affected by its own maximum error value. In an actual application there will be facts and rules which are no less than true, facts and rules which are reliable up to a degree of confidence of, say, 95 percent, and so on.

Let us consider the simplified situation. We could take advantage of the fact that in a p -valued EP system we actually have p subsystems $(\bigvee_m, \bigwedge_m [\cdot], m \in I(p))$, each isomorphic to a classical p -valued Post system, and all interleaved in one and the same EP system. Since it is isomorphic to a Post system, each of these subsystems is functionally complete. Thus, we could single out the proper one of these subsystems and work just with it, ignoring the rest of the full EP system. If, e.g., $p = 100$, and the available data are all reliable up to a degree of confidence of 80 percent, we could single out the subsystem based on $(\bigvee_m, \bigwedge_m [\cdot], m = 80)$. The problem with this choice is that the cycle operation $[\cdot]$ does not behave as a negation (though, in his early work, Post treats it as a generalized negation).

As we shall see, a better choice is to adopt a subsystem based on the operations $\sim, \bigvee_m, \bigwedge_m, \bigvee_{p-m}, \bigwedge_{p-m}$. Note that, by Proposition 7, the operations $\bigvee_{p-m}, \bigwedge_{p-m}$ are actually unessential, since they can be expressed in terms of $\sim, \bigvee_m, \bigwedge_m$. Again, m is to be selected according to the desired degree of approximation. We denote $EP(p, m)$ such a subsystem of $EP(p)$.

The simplified situation considered above cannot be expected to occur very frequently in practice. In real

applications we must be able to derive consequences from facts and rules given with different degrees of truth. As we shall see, each single derivation step is performed in a well-defined $EP(p, m)$ subsystem. However, all the $EP(p, m)$ subsystems, $m > \frac{1}{2}$, are in general used in a full derivation process.

In the following, we single out a fixed $EP(p, m)$ subsystem and investigate it from a logical, rather than an algebraic, point of view. Before doing that, however, we state the two basic inference rules of an $EP(p)$ system. In a rather evident way, these are straight generalizations of *modus ponens* and *generalization* of ordinary two-valued logic.

Definition 8

A well-formed formula S is m -true (under an interpretation I) if and only if $T(S) \geq u(m)$ (under I).

Remark Obvious references to an interpretation are omitted.

Definition 9

A WFF S is m -false iff $T(S) \leq 1 - u(m)$.

The two basic inference rules of a p -valued EP system can now be formulated as follows. For every pair of WFFs F, G , and all m in $I(p)$ such that $u(m) > \frac{1}{2}$,

- If F and $F \supset_m G$ are m -true, infer that G is m -true.
- If F is m -true, infer that $(\bigwedge_m x)F$ is m -true.

The justification for so straight a generalization of two-valued logic can be made clear by examining a few immediate properties of any $EP(p, m)$ subsystem. In what follows, m is such that $u(m) > \frac{1}{2}$.

Proposition 10

A WFF S is m -true iff $\sim S$ is m -false.

Proof $T(S) \geq u(m)$ iff $-T(S) \leq -u(m)$ iff $1 - T(S) \leq 1 - u(m)$. Thus, by RE1, $T(S) \geq u(m)$ iff $T(\sim S) \leq 1 - u(m)$. By Definition 9, the proposition follows.

Proposition 11

$A \bigvee_m B$ is m -true iff either A or B is m -true, or both are m -true.

Proof Let A be m -true, i.e., $T(A) \geq u(m)$. If $T(B) \geq u(m)$, by RE3 we have $T(A \bigvee_m B) = \min(TA, TB) \geq u(m)$; hence $A \bigvee_m B$ is m -true. If $T(B) < u(m)$, by RE3 we have $T(A \bigvee_m B) = \max(TA, TB) = TA \geq u(m)$; hence $A \bigvee_m B$ is again m -true. Since $\bigvee_m$ is commutative, the *if* part of the proposition is proved. Let $A \bigvee_m B$ be m -true, i.e., $T(A \bigvee_m B) \geq u(m)$. By RE3, either $T(A \bigvee_m B) = \min(TA, TB)$ or $T(A \bigvee_m B) = \max(TA, TB)$. In the former case, $T(A), T(B) \geq u(m)$ and both A and B are m -true. In the latter case, since $T(A \bigvee_m B) \geq u(m)$, we have either $T(A) \geq u(m)$ or $T(B) \geq u(m)$.

Proposition 12

$A \overset{m}{\wedge} B$ is m -true if both A and B are m -true.

Proof Analogous to that of Proposition 11.

Proposition 13

If A is either m -true or m -false, then $A \overset{m}{\supset} A$ is m -true.

Proof Let A be m -true. By Proposition 10, $\sim A$ is m -false. Then, by Proposition 11, $(\sim A) \overset{m}{\vee} A$ is m -true. Thus, by RE5, $A \overset{m}{\supset} A$ is m -true. If A is m -false, then $\sim A$ is m -true, and $A \overset{m}{\supset} A$ is again m -true.

Proposition 14

Assume that A, B, C are each either m -true or m -false. If $A \overset{m}{\supset} B$ and $B \overset{m}{\supset} C$ are both m -true, then $A \overset{m}{\supset} C$ is m -true.

Proof By RE5 and Proposition 2,

$$\begin{aligned} H &= (A \overset{m}{\supset} B) \overset{m}{\wedge} (B \overset{m}{\supset} C) \\ &\approx (\sim A \overset{m}{\vee} B) \overset{m}{\wedge} (\sim B \overset{m}{\vee} C) \\ &\approx (\sim A \overset{m}{\wedge} \sim B) \overset{m}{\vee} (\sim A \overset{m}{\wedge} C) \overset{m}{\vee} (B \overset{m}{\wedge} \sim B) \overset{m}{\vee} (B \overset{m}{\wedge} C). \end{aligned}$$

By Propositions 10 and 12, $(\sim B) \overset{m}{\wedge} B$ is m -false. Let $A \overset{m}{\supset} B$ and $B \overset{m}{\supset} C$ be m -true. By Proposition 12, H is m -true. (1) Assume A is m -true. By Propositions 10 and 12, both $(\sim A) \overset{m}{\wedge} (\sim B)$ and $(\sim A) \overset{m}{\wedge} C$ are m -false. Since H is m -true, by Proposition 11, $B \overset{m}{\wedge} C$ is m -true. By Proposition 12, that means that both B and C are m -true. Thus $(\sim A) \overset{m}{\vee} C$, i.e., $A \overset{m}{\supset} C$, is m -true. (2) Assume A is m -false. Then $(\sim A) \overset{m}{\vee} C$, hence $A \overset{m}{\supset} C$, is m -true by Propositions 10 and 11.

Proposition 15

If both A and B are either m -true or m -false, then $(A \overset{m}{\wedge} B) \overset{m}{\supset} A$ is always m -true.

Proof $A \overset{m}{\wedge} B$ is either m -true or m -false. By RE5, $(A \overset{m}{\wedge} B) \overset{m}{\supset} A$ is logically equivalent to $\sim(A \overset{m}{\wedge} B) \overset{m}{\vee} A$. If $A \overset{m}{\wedge} B$ is m -true, A is m -true by Proposition 12; thus, by Proposition 11, $\sim(A \overset{m}{\wedge} B) \overset{m}{\vee} A$ is m -true. If $A \overset{m}{\wedge} B$ is m -false, then $\sim(A \overset{m}{\wedge} B)$ is m -true, whence $\sim(A \overset{m}{\wedge} B) \overset{m}{\vee} A$ is again m -true.

Proposition 16

If A is either m -true or m -false, then $A \overset{m}{\supset} (A \overset{m}{\vee} B)$ is m -true.

Proof By RE5, $A \overset{m}{\supset} (A \overset{m}{\vee} B)$ is logically equivalent to $\sim A \overset{m}{\vee} (A \overset{m}{\vee} B)$, whence the conclusion, by Propositions 1, 10, and 11.

Proposition 17

$A \overset{m}{\supset} B$ is logically equivalent to $\sim B \overset{m}{\supset} \sim A$.

Proof By RE5, $A \overset{m}{\supset} B \approx (\sim A) \overset{m}{\vee} B$. Again by RE5, $\sim B \overset{m}{\supset} \sim A \approx (\sim \sim B) \overset{m}{\vee} \sim A \approx B \overset{m}{\vee} \sim A \approx (\sim A) \overset{m}{\vee} B$.

Propositions 10–17 suggest that each $EP(p, m)$ subsystem exhibits formal properties similar to those of ordinary two-valued logic. It is thus natural to ask whether these systems have properties which make them suitable for incorporation into problem-solving systems.

Inference rules other than *modus ponens* and *generalization* are used to derive theorems in two-valued logic. Various theorem-proving techniques are based on the inference rule known as the *resolution principle* [30]. This rule of inference permits a new clause to be derived from two given clauses. The derived clause is satisfiable (i.e., has a model) whenever the two parent clauses are satisfiable.

Resolution applies to WFFs in clausal form. A clause is defined as a WFF consisting of a disjunction of literals. A literal is an atomic formula or the negation of an atomic formula (i.e., a formula with a $\sim$ in front of it). When resolution is used as a rule of inference, the set of WFFs representing the available knowledge base is first put into clausal form, i.e., is converted into a conjunction of clauses where all variables are universally quantified. Any WFF in the language of first-order two-valued logic can be so converted; thus, clauses are a general form in which to express WFFs.

Techniques based on the resolution principle play a major role in automated reasoning. Most theorem-proving and logic programming systems rely upon inference methods which are variants of the resolution principle. Thus, understanding whether a system of logic is suited for incorporation into actual problem-solving systems is almost equivalent to investigating whether resolution is applicable in it.

It turns out that the resolution principle is applicable as an inference rule in each $EP(p, m)$ subsystem. Also, each $EP(p, m)$ subsystem is complete to resolution. That means that, given any pair F, G of WFFs in one of these subsystems, if G is m -true in all interpretations in which F is m -true, then G can be derived from F by means of a resolution-refutation procedure.

These results follow from the results of Chang and Lee [19] and Lee [20] on the applicability of the resolution principle to fuzzy logic, and the basic lemma proved in Section 4. We discuss briefly how the results of Chang and Lee apply in our case. A detailed proof can be found in [4].

First, we note that, for $m = 0$, the $EP(p, m)$ subsystem reduces to fuzzy logic. Indeed, the only difference between the two systems consists in the truth-set they adopt: The $EP(p, 0)$ system uses the set $V(p)$, while fuzzy logic uses the unit interval $[0, 1]$. However, the properties of the truth-set actually used in fuzzy logic are certain algebraic properties that are common to $V(p)$ and $[0, 1]$. That is, the truth-set is assumed to be a complete distributive lattice with 0, 1 whose associated ordering is total. Both $V(p)$ and the unit interval

have these properties. The continuity properties of the unit interval are never actually used in the construction of fuzzy logic. In view of that, we henceforth make no distinction between fuzzy logic and the $EP(p, 0)$ subsystem.

We now recall the results obtained by Chang and Lee [19] and Lee [20] regarding fuzzy logic.

Let F, G be any WFFs in the language of fuzzy logic. An interpretation I satisfies F if $T(F) \geq \frac{1}{2}$ under I . I falsifies F if $T(F) < \frac{1}{2}$ under I . F is unsatisfiable if it is falsified by every interpretation. G is a logical consequence of F if $F \wedge \sim G$ is unsatisfiable. Equivalently, G is a logical consequence of F if G is satisfied in all interpretations in which F is satisfied.

Chang and Lee proved that, given any set S of WFFs in the language of fuzzy logic, if all the clauses in S have truth values in the interval $[a, b]$, where $\frac{1}{2} < a \leq b \leq 1$, then we are guaranteed that all the logical consequences obtained by repeatedly applying the resolution principle will have truth value between a and b .

Now let us consider a formula A in the $EP(p, 0)$ subsystem. According to our definitions, A is $\frac{1}{2}$ -true under an interpretation I if and only if I satisfies A in the sense of fuzzy logic. A is $\frac{1}{2}$ -false under I if and only if I falsifies A . Taking that into account, and remembering that $EP(p, m)$ is isomorphic to $EP(p, 0)$ by virtue of $[.]^m$, the above results can be reformulated as follows. If every clause in a set of clauses $EP(p, m)$ has truth value greater than or equal to $u(m)$, and the most unreliable clause has truth value a while the most reliable clause has truth value b , then we are guaranteed that any clause obtained by the resolution principle has truth value in the range $[a, b]$.

In conclusion, inference in a p -valued Extended Post system can be made as follows. Assume that we start with a set S of axiomatic ground clauses with truth values in $V(p)$. Let S' be a subset of S , and assume that there exists an m in $V(p)$, $m > \frac{1}{2}$, such that all the clauses in S' are either m -true or m -false. Then we may invoke the $EP(p, m)$ subsystem to derive logical consequences from S' in the usual manner, e.g., by Robinson's resolution rule. All the conclusions thus derived will be m -true.

In the rest of this section we discuss some examples.

Example 1 Let $p = 11$; thus $V(p) = \{0, 0.1, \dots, 0.9, 1\}$. Let $Ax [A(x)]$, if there is danger of confusion mean "x is an athlete," Ix mean "x is intelligent," Cx mean "x has courage," Sx mean "x will succeed in his career," Hx mean "x will be happy." Assume the axioms

$$Ax \supset_{\frac{8}{8}} Cx, \quad (A1)$$

$$Ix \wedge_{\frac{6}{6}} Cx \supset_{\frac{7}{7}} Sx. \quad (A2)$$

Let a be an individual. Suppose that we know the truth values of the clauses Aa and Ia . What consequences can be derived from A1, A2?

If $T(Aa) \geq u(8) = 0.8$, we are able to infer the inequality $T(Ca) \geq 0.8$ from A1. If, besides, $T(Ia) \geq u(6) = 0.6$, then $T(Ia \wedge_{\frac{6}{6}} Ca) \geq 0.8 > 0.7 = u(7)$. Thus, we are able to infer $T(Sa) \geq 0.7$ from axiom A2.

If $T(Aa) \geq 0.8$ and $T(Ia) < 0.6$, then again we are able to infer $T(Ca) \geq 0.8$. However, $Ia \wedge_{\frac{6}{6}} Ca$ evaluates to less than 0.6, hence less than 0.7; thus we cannot make any inference about the career of a .

If $T(Aa) < 0.8$, then we are not able to infer anything from A1, A2.

Example 2 Let all the predicate symbols preserve their meanings. According to axiom A1, as soon as $T(Aa) \geq 0.8$, we infer $T(Ca) \geq 0.8$. This may seem too drastic in some situations. One might wish a weaker rule, e.g., one allowing $T(Ca) \geq 0.6$ to be inferred from $T(Aa) \geq 0.8$. Then, we replace axiom A1 with

$$Ax \supset_{\frac{8}{8}} (Cx \wedge_{\frac{6}{6}} K(0.8)). \quad (B1)$$

Here $K(0.8)$ denotes an 8-tautology (i.e., according to the definition given in Section 4, a formula which takes the value 0.8 under any interpretation). Again, if $T(Aa) < 0.8$, we are not able to infer anything. If $T(Aa) \geq 0.8$, then $T(Ca \wedge_{\frac{6}{6}} K(0.8)) \geq 0.8$, whence $T(Ca) \geq 0.6$. Now, if $T(Ia) \geq 0.7$, from axiom A2 we infer that $T(Sa) \geq 0.7$, since in this case $T(Ia \wedge_{\frac{6}{6}} Ca) = \max\{T(Ia), T(Ca)\} = T(Ia)$. If $T(Ia) < 0.7$, we cannot make any inference about the career of a .

Example 3 Let all the predicate symbols preserve their meanings. We keep axiom B1 and replace axiom A2 with

$$(Ix \wedge_{\frac{6}{6}} Cx) \supset_{\frac{7}{7}} (Sx \wedge_{\frac{5}{5}} Hx). \quad (B2)$$

If $T(Aa) < 0.8$, we infer nothing. If $T(Aa) \geq 0.8$ and $T(Ia) < 0.7$, we infer only that $T(Ca) \geq 0.6$. If $T(Aa) \geq 0.8$ and $T(Ia) \geq 0.7$, then we infer that $T(Sa \wedge_{\frac{5}{5}} Ha) \geq 0.7$; that is, both Sa and Ha are 5-true, and at least one of them is 7-true.

6. Remarks on the use of EP logic in rule-based systems

The examples at the end of the last section show how reasoning can be done in EP systems. It was seen that reasoning is possible provided certain numerical thresholds are specified. Thus, in axiom A1 there is one threshold value associated with the implication sign; in axiom A2 two thresholds occur, one associated with the conjunction in the antecedent and the other associated with the implication; analogously for the other axioms.

In view of an actual use of EP logic in rule-based systems, we must show, or at least indicate, how these numerical values can be determined. For example, with reference to axiom A1, how do we know that $Ax \supset_{\frac{8}{8}} Cx$ and that it is not,

say, $Ax \supset Cx$? Indeed, there is no unique translation of a sentence like "the athletes are courageous" into the *EP* language (by the way, axioms A1 and A2 are two distinct translations).

The answer to this question is suggested by the nature (from a logical point of view) of such sentences as "the athletes are courageous." Sentences like this are factual (or logically indeterminate, as logicians also say). Their truth value cannot be established on purely logical grounds, on the mere basis of their form. It is necessary to institute observations. Thus, only the observation of the actual world can tell us whether the degree of truth of the statement asserting that the athletes are courageous is 0.8 or 0.6.

Pursuing this line of reasoning, we must specify a methodology for assessing the numerical values in the axioms from our knowledge of the world. One of the most known and theoretically motivated methodologies is the statistical approach. We discuss its application in the present context in some detail. Its strict application would be impractical in most cases; however, it provides a general criterion for the selection of numerical values in the axioms.

Let us first examine a specially suggestive example. Let Px stand for " x took examination P ," and Qx stand for " x took examination Q ." In many-valued logic we may assume that $T(Px) \geq p$ means that x took examination P and gained p marks, and analogously for $T(Qx) \geq p$. Let us analyze the implication

$$Px \supset (Qx \wedge K(0.8)), \quad (6)$$

which is formally identical to axiom B1 in Example 2. On the basis of the theory developed before, the inclusion of this axiom in an axiom system is equivalent to the assumption that $T(Px) \geq 0.8$ implies $T(Qx) \geq 0.6$. In other words, if we accept this axiom, it means that we expect that an individual taking examination P and gaining 8 marks or more would gain no less than 6 marks if he were to take examination Q .

Now the question is this: What factual evidence can lead us to make assumption (6)? The answer is very simple. We accept this rule if sufficient data are available for us to infer that almost surely (i.e., with probability 1), if an individual gains no less than 8 marks in examination P , then he or she will be able to gain at least 6 marks in examination Q . In other words, the conditional probability of the event $T(Qx) \geq 0.6$ under the condition $T(Px) \geq 0.8$ should be equal to 1:

$$\text{prob}\{T(Qx) \geq 0.6 \mid T(Px) \geq 0.8\} = 1.$$

Note that we are in a Bayesian setting. Bayes' theorem holds that

$$\begin{aligned} \text{prob}\{T(Qx) \geq 0.6 \mid T(Px) \geq 0.8\} \text{prob}\{T(Px) \geq 0.8\} \\ = \text{prob}\{T(Px) \geq 0.8 \mid T(Qx) \geq 0.6\} \text{prob}\{T(Qx) \geq 0.6\}. \end{aligned}$$

Note that, since $\text{prob}\{T(Px) \geq 0.8 \mid T(Qx) \geq 0.6\} \leq 1$, we have

$$\text{prob}\{T(Px) \geq 0.8\} \leq \text{prob}\{T(Qx) \geq 0.6\}.$$

This is obvious: In this case the event $T(Px) \geq 0.8$ almost surely implies the event $T(Qx) \geq 0.6$; thus, the probability of the event $T(Px) \geq 0.8$ cannot be greater than the probability of $T(Qx) \geq 0.6$.

Now we try to rephrase these conclusions in terms of axiom B1 in Example 2 above. This axiom has exactly the same form as the one just studied, but the meaning is different: Ax stands for " x is an athlete," and Cx stands for " x has courage." We can think of $T(Ax)$ and $T(Cx)$ as measures of the extent to which x is an athlete and, respectively, a courageous individual. Then, if in the actual world the individuals who are athletes to an extent no less than 0.8 happen, with probability 1, to be courageous to an extent no less than 0.6, we will accept axiom B1.

Let us now consider axiom B2 in Example 3. Although this axiom is more complex than B1, nothing changes in principle, and we have only to rephrase the conclusions regarding B1 in this more involved situation. In other words, the inclusion of B2 in the knowledge base is equivalent to the assumption that $T(Ix \wedge Cx) \geq 0.7$ almost surely implies $T(Sx \wedge Hx) \geq 0.7$. The first inequality is equivalent to the set of inequalities

$$T(Ix) \geq 0.6, \quad (7)$$

$$T(Cx) \geq 0.6, \quad (8)$$

$$T(Ix) \geq 0.7 \quad \text{or} \quad T(Cx) \geq 0.7. \quad (9)$$

The second inequality is equivalent to the set of inequalities

$$T(Sx) \geq 0.5, \quad (10)$$

$$T(Hx) \geq 0.5, \quad (11)$$

$$T(Sx) \geq 0.7 \quad \text{or} \quad T(Hx) \geq 0.7. \quad (12)$$

Thus, the condition that must be satisfied in order that B2 may be accepted as a rule is the following: There must be empirical evidence that, whenever (7)–(9) are satisfied, then (10)–(12) are also satisfied. More specifically, the conditional probability of the event represented by the inequalities (10)–(12), given that the event represented by (7)–(8) is true, should be equal to 1.

The statistical approach has sound theoretical motivations, but in most cases it cannot be put into practice too strictly. How can we estimate the fraction of the athletes who are courageous? Shall we draw a random sample of athletes and perform some tests to assign an individual degree of courage to each member of the sample?

It is considered more practical to rely upon the degree of confidence of the expert in his or her expressed rule. Thus, in the case of *EP* logic, the expert is expected to provide the rules, together with the thresholds that occur in them. The

statistical methodology is of use in that it provides the expert with a criterion for assessing the thresholds. For example, an expert will supply rule B2 if he or she feels confident that, whenever (7)–(9) are satisfied, then almost surely (10)–(12) are also satisfied.

7. Concluding remarks

In this paper we have presented a system of many-valued logic suited for handling approximate or imprecise knowledge.

On the one hand, the system has an algebraic foundation, and on the other hand, it provides for translation of certain recurrent patterns of informal reasoning to logical form.

A well-known requirement set forth by Birkhoff [22] on many-valued logics, that the truth assignment function (denoted T in this paper) should be a morphism from the set of all the propositions to the set of the possible truth values, is satisfied in this system. Indeed, the set of all the propositions and the truth-set are here equipped with isomorphic algebraic structures, based on the two unary operations $\sim$ and $[\cdot]$, a set of generalized ANDs, and a conjugate set of generalized ORs.

In the system of logic presented in this paper, inference can be performed by Robinson's resolution principle, as in ordinary two-valued logic. Thus the system is complete to resolution, thereby satisfying one basic prerequisite for incorporation into rule-based systems.

References

1. N. J. Nilsson, "Probabilistic Logic," *Artif. Intell.* **28**, 71–87 (1986).
2. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.
3. W. F. Eddy and G. P. Pei, "Structures of Rule-Based Belief Functions," *IBM J. Res. Develop.* **30**, 93–101 (1986).
4. S. Di Zeno, "A New Many-Valued Logic and Its Application to Approximate Reasoning," *Proceedings of the IFIP 10th World Computer Congress*, Dublin, September 1–5, 1986, Elsevier Science Publishers, Amsterdam, pp. 421–427.
5. S. Di Zeno, "Multiple Boolean Algebras and Their Application to Fuzzy Sets," *Info. Sci.* **35**, 111–132 (1985).
6. J. Lukasiewicz, "O logice trojwartosciowej," *Ruc. Filozof.* **5**, 169–171 (1920).
7. E. L. Post, "Introduction to a General Theory of Elementary Propositions," *Amer. J. Math.* **43**, 163–185 (1921).
8. J. B. Rosser and A. R. Turquette, *Many-Valued Logics*, North-Holland Publishing Co., Amsterdam, 1952.
9. R. Ackermann, *An Introduction to Many-Valued Logics*, Routledge and Kegan, New York, 1971.
10. N. Rescher, *Many-Valued Logics*, McGraw-Hill Book Co., Inc., New York, 1969.
11. R. G. Wolf, "A Survey of Many-Valued Logic (1966–1974)," *Modern Uses of Multiple Valued Logic*, J. Dunn and G. Epstein, Eds., D. Reidel Publishing Co., Dordrecht, Netherlands, 1977, pp. 167–323.
12. S. L. Hurst, "Multiple-Valued Logic—Its Status and Its Future," *IEEE Trans. Computers* **C-33**, 1160–1179 (1984).
13. S. Guccione, S. Termini, and R. Tortora, "In the Labyrinth of Many-Valued Logics," *Proceedings of the 13th International Symposium on Multiple-Valued Logic*, Kyoto, Japan, 1983, pp. 47–53.
14. L. A. Zadeh, "Fuzzy Sets," *Info. & Control* **8**, 338–353 (1965).
15. R. E. Bellman and L. A. Zadeh, "Local and Fuzzy Logics," *Modern Uses of Multiple Valued Logic*, J. Dunn and G. Epstein, Eds., D. Reidel Publishing Co., Dordrecht, Netherlands, 1977, pp. 105–165.
16. L. A. Zadeh, "A Theory of Approximate Reasoning," *Machine Intelligence*, Vol. 9, J. E. Hayes, D. Michie, and L. I. Mikulich, Eds., Elsevier, New York, 1979, pp. 149–194.
17. H. Prade, "A Computational Approach to Approximate and Plausible Reasoning with Applications to Expert Systems," *IEEE Trans. Pattern Anal. & Mach. Intell.* **PAMI-7**, 260–283 (1985).
18. P. N. Marinos, "Fuzzy Logic and Its Applications to Switching Systems," *IEEE Trans. Computers* **C-18**, 343–348 (1969).
19. C. L. Chang and R. C. T. Lee, "Some Properties of Fuzzy Logic," *Info. & Control* **19**, 417–431 (1971).
20. R. C. T. Lee, "Fuzzy Logic and the Resolution Principle," *J. ACM* **19**, 109–119 (1972).
21. A. R. Aronson, B. E. Jacobs, and J. Minker, "A Note on Fuzzy Deduction," *J. ACM* **27**, 599–603 (1980).
22. G. Birkhoff, *Lattice Theory*, 3rd Ed., American Mathematical Society College Publications, Providence, RI, 1967.
23. P. C. Rosenbloom, "Post Algebras. I. Postulates and General Theory," *Amer. J. Math.* **64**, 167–188 (1942).
24. G. Epstein, "The Lattice Theory of Post Algebras," *Trans. Amer. Math. Soc.* **95**, 63–74 (1960).
25. Ph. Dwingier, "A Survey of the Theory of Post Algebras and Their Generalizations," *Modern Uses of Multiple Valued Logic*, J. Dunn and G. Epstein, Eds., D. Reidel Publishing Co., Dordrecht, Netherlands, 1977, pp. 53–75.
26. H. Rasiowa, *An Algebraic Approach to Non-Classical Logics*, North-Holland Publishing Co., Amsterdam, 1974.
27. H. Rasiowa, "Many-Valued Algorithmic Logic as a Tool to Investigate Programs," *Modern Uses of Multiple Valued Logic*, J. Dunn and G. Epstein, Eds., D. Reidel Publishing Co., Dordrecht, Netherlands, 1977, pp. 79–102.
28. E. Orłowska, "The Resolution Principle for Omega-Valued Logic," *Fund. Informat.* **II**, 1–15 (1978).
29. E. Orłowska, "Mechanical Proof Methods for Post Logics," *Logique et Analyse* **110**, 173–192 (1985).
30. J. A. Robinson, "A Machine-Oriented Logic Based on the Resolution Principle," *J. ACM* **12**, 23–41 (1965).

Received May 28, 1987; accepted for publication February 5, 1988

Silvano Di Zeno IBM Italy, Rome Scientific Center, via Giorgione, 159, 00147 Rome, Italy. Dr. Di Zeno received the electrical engineering degree from the University of Genoa in 1962. He joined IBM in 1964; from 1965 to 1970 he was also an assistant professor of mathematics at the University of Genoa. From 1976 to 1979 he was program manager of an image processing industry project, and in 1980 he joined the IBM Rome Scientific Center, where he is currently responsible for research projects in image recognition and understanding, robot vision, and automated reasoning. Dr. Di Zeno is a member of the Italian chapter of the European Association for Theoretical Computer Science, the Italian Association for Automated Computing, the Italian Association for Pure and Applied Biophysics, and the editorial board of *Image and Vision Computing*.